CSE- 1006 LAB Assignment 3.1

Academic year: 2021-2022 **Semester:** WIN

Faculty Name: Dr. Arun kumar Gopu **Date:** 17/3/2022

Student name: M.Taran **Reg. no.:** 19BCE7346

Data Recoding:

Data Recoding in R can be done by either replacing data in an existing field or recoding into a new field based on criteria you specify

EXERCISES

Consider a numeric vector x <- c(3,4,5,6,7,8)

```
>
> x <- c(3,4,5,6,7,8)
>
```

Write a command to recode the values less than 6 with zero in the vector x

x[x<6]=0

```
> x[x<6]=0
> x
[1] 0 0 0 6 7 8
```

Write a command to recode the values between 4 and 8 with 100

x[x>6 & x<8]=100

```
> x[x>6 & x<8]=100
>
> x
[1]   0   0   0   6 100   8
```

Write a command to recode the values that are less than 5 or greater than 6 with 50

x[x<5 | x>6]<-50

```
> x[x<5 | x>6]<-50
> x
[1] 50 50 50  6 50 50
```

Write a command to recode the values less than 6 with NA in the vector x

x[x<6]<-NA

```
> x[x<6]<-NA
> x
[1] 50 50 50  6 50 50
```

Write a command to recode the values between 4 and 8 with NA

x[x>4 & x<8] <-NA

```
> x[x>4 & x<8] <-NA
> x
[1] 50 50 50 NA 50 50
```

Write a command to recode the values that are less than 5 or greater than 6 with NA

x[x<5 | x>6]<-NA

```
> x[x<5 | x>6]<-NA
> x
[1] NA NA NA NA NA NA
```

Count number of NA values after each operation

sum(is.na(x))

```
> sum(is.na(x))
[1] 6
```

Find mean of x (Hint: exclude NA values)

mean(x, na.rm=TRUE)

```
> mean(x, na.rm=TRUE)
[1] NaN
```

Find median of x (Hint: exclude NA values)

median(x, na.rm=TRUE)

```
> median(x, na.rm=TRUE)
[1] NA
```

Write a command to recode the values less than 6 with “NA” (enclose NA with double quotes) in the vector x

x=as.integer(x)

```
> x[x<6] <- "NA"
> x
[1] "NA" "NA" "NA" "6"  "7"  "8"
```

Write a command to recode the values between 4 and 8 with “NA”

```
> x[x>4 & x<8] <- "NA"
> x
[1] "3"  "4"  "NA" "NA" "NA" "8"
```

Write a command to recode the values that are less than 5 or greater than 6 with “NA”

```
> x <- c(3,4,5,6,7,8)
> x[x<5 | x>6]<- "NA"
> x
[1] "NA" "NA" "5"  "6"  "NA" "NA"
```

Count number of NA values after each operation

sum(is.na(x)) will not consider "NA" as it is a string

Find mean of x (Hint: exclude NA values)

Mean cannot be calculated as "NA" is a string

Find median of x (Hint: exclude NA values)

Median cannot be calculated as "NA" is a string

What is the difference between NA and “NA”

NA is not available (missing value) whereas "NA" is a string

EXERCISES

Consider the given vectors:

A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)

B <- c(3, 2, NA, 5, 3, 7, NA, “NA”, 5, 2, 6)

```
> A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
> B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)
```

Find the length of the vector A

```
> length(A)
[1] 11
```

Find the length of the vector B

```
> length(B)
[1] 11
```

Sort the values in vector A and put it in p (Hint: use function sort())

```
> p <- sort(A)
> p
[1] 2 2 3 3 5 5 6 7
```

Find the length of p

```
> length(p)
[1] 8
```

Sort the values in vector B and put it in q

```
> q <- sort(B)
> q
[1] "2"  "2"  "3"  "3"  "5"  "5"  "6"  "7"  "NA"
```

Find the length of q

```
> length(q)
[1] 9
```

What did you infer from the above results

EXERCISES

Let us work on dataset – airquality

```
> data(airquality)
```

Print the dataset airquality

```
> print(airquality)
    Ozone Solar.R Wind Temp Month Day
1      41     190  7.4   67     5   1
2      36     118  8.0   72     5   2
3      12     149 12.6   74     5   3
4      18     313 11.5   62     5   4
5      NA      NA 14.3   56     5   5
6      28      NA 14.9   66     5   6
7      23     299  8.6   65     5   7
8      19      99 13.8   59     5   8
9       8      19 20.1   61     5   9
10     NA     194  8.6   69     5  10
11      7      NA  6.9   74     5  11
12     16     256  9.7   69     5  12
13     11     290  9.2   66     5  13
14     14     274 10.9   68     5  14
15     18      65 13.2   58     5  15
16     14     334 11.5   64     5  16
17     34     307 12.0   66     5  17
18      6      78 18.4   57     5  18
19     30     322 11.5   68     5  19
20     11      44  9.7   62     5  20
21      1       8  9.7   59     5  21
22     11     320 16.6   73     5  22
23      4      25  9.7   61     5  23
24     32      92 12.0   61     5  24
25     NA      66 16.6   57     5  25
26     NA     266 14.9   58     5  26
27     NA      NA  8.0   57     5  27
```

Print the structure of the dataset airquality

```
> str(airquality)
'data.frame':   153 obs. of  6 variables:
 $ Ozone  : int  41 36 12 18 NA 28 23 19 8 NA ...
 $ Solar.R: int  190 118 149 313 NA NA 299 99 19 194 ...
 $ Wind   : num  7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.
6 ...
 $ Temp   : int  67 72 74 62 56 66 65 59 61 69 ...
 $ Month  : int  5 5 5 5 5 5 5 5 5 5 ...
 $ Day    : int  1 2 3 4 5 6 7 8 9 10 ...
```

Print the summary of all the variables of the dataset airquality (Hint: Use function summary())

```
> summary(airquality)
     Ozone           Solar.R          Wind
 Min.   :  1.00   Min.   :  7.0   Min.   : 1.700
 1st Qu.: 18.00   1st Qu.:115.8   1st Qu.: 7.400
 Median : 31.50   Median :205.0   Median : 9.700
 Mean   : 42.13   Mean   :185.9   Mean   : 9.958
 3rd Qu.: 63.25   3rd Qu.:258.8   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
 NA's   :37       NA's   :7
      Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0
```

How many of the variables (columns) are in the dataset airquality

```
> ncol(airquality)
[1] 6
```

```
> colSums(is.na(airquality))
  Ozone Solar.R    Wind    Temp   Month     Day
     37       7       0       0       0       0
```

How many observations (rows) are in the dataset airquality

```
> nrow(airquality)
[1] 153
```

Observe the results of summary() function on dataset airquality. Attributes Ozone and Solar.R have missing values. Number of missing values are displayed at the bottom of each column if any.

What are the values getting displayed when we use summary() function

What is quartile how to find them

What are 1st and 3rd quartiles

Copy the dataset airquality to aq (Better work on a copy of original data instead of working on original data to avoid the loss of information)

```
> aq <- airquality
>
```

Print the dataset aq

```
> aq
    Ozone Solar.R Wind Temp Month Day
1      41     190  7.4   67     5   1
2      36     118  8.0   72     5   2
3      12     149 12.6   74     5   3
4      18     313 11.5   62     5   4
5      NA      NA 14.3   56     5   5
6      28      NA 14.9   66     5   6
7      23     299  8.6   65     5   7
8      19      99 13.8   59     5   8
9       8      19 20.1   61     5   9
10     NA     194  8.6   69     5  10
11      7      NA  6.9   74     5  11
12     16     256  9.7   69     5  12
13     11     290  9.2   66     5  13
14     14     274 10.9   68     5  14
15     18      65 13.2   58     5  15
16     14     334 11.5   64     5  16
17     34     307 12.0   66     5  17
18      6      78 18.4   57     5  18
```

```
> print(aq)
    Ozone Solar.R Wind Temp Month Day
1      41     190  7.4   67     5   1
2      36     118  8.0   72     5   2
3      12     149 12.6   74     5   3
4      18     313 11.5   62     5   4
5      NA      NA 14.3   56     5   5
6      28      NA 14.9   66     5   6
7      23     299  8.6   65     5   7
8      19      99 13.8   59     5   8
9       8      19 20.1   61     5   9
10     NA     194  8.6   69     5  10
11      7      NA  6.9   74     5  11
12     16     256  9.7   69     5  12
13     11     290  9.2   66     5  13
14     14     274 10.9   68     5  14
15     18      65 13.2   58     5  15
16     14     334 11.5   64     5  16
17     34     307 12.0   66     5  17
18      6      78 18.4   57     5  18
```

Print the structure of the dataset aq

```
> str(aq)
'data.frame':   153 obs. of  6 variables:
 $ Ozone  : int  41 36 12 18 NA 28 23 19 8 NA ...
 $ Solar.R: int  190 118 149 313 NA NA 299 99 19 194 ...
 $ Wind   : num  7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.
6 ...
 $ Temp   : int  67 72 74 62 56 66 65 59 61 69 ...
 $ Month  : int  5 5 5 5 5 5 5 5 5 5 ...
 $ Day    : int  1 2 3 4 5 6 7 8 9 10 ...
```

Print the summary of all the variables of the dataset aq (Hint: Use function summary())

```
> summary(aq)
     Ozone           Solar.R          Wind       
 Min.   :  1.00   Min.   :  7.0   Min.   : 1.700  
 1st Qu.: 18.00   1st Qu.:115.8   1st Qu.: 7.400  
 Median : 31.50   Median :205.0   Median : 9.700  
 Mean   : 42.13   Mean   :185.9   Mean   : 9.958  
 3rd Qu.: 63.25   3rd Qu.:258.8   3rd Qu.:11.500  
 Max.   :168.00   Max.   :334.0   Max.   :20.700  
 NA's   :37       NA's   :7                       
      Temp           Month            Day       
 Min.   :56.00   Min.   :5.000   Min.   : 1.0  
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0  
 Median :79.00   Median :7.000   Median :16.0  
 Mean   :77.88   Mean   :6.993   Mean   :15.8  
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0  
 Max.   :97.00   Max.   :9.000   Max.   :31.0  
```

Print top 6 observations

```
> head(aq)
  Ozone Solar.R Wind Temp Month Day
1    41     190  7.4   67     5   1
2    36     118  8.0   72     5   2
3    12     149 12.6   74     5   3
4    18     313 11.5   62     5   4
5    NA      NA 14.3   56     5   5
6    28      NA 14.9   66     5   6
```

Print last 6 observations

```
> tail(aq)
    Ozone Solar.R Wind Temp Month Day
148    14      20 16.6   63     9  25
149    30     193  6.9   70     9  26
150    NA     145 13.2   77     9  27
151    14     191 14.3   75     9  28
152    18     131  8.0   76     9  29
153    20     223 11.5   68     9  30
```

Replace the NA values in the attribute Ozone in aq by zero

```
> aq[is.na(aq)] <- 0
```

Print the summary of all the variables of the dataset aq

```
> summary(aq)
     Ozone           Solar.R          Wind
 Min.   :  0.00   Min.   :  0.0   Min.   : 1.700
 1st Qu.:  4.00   1st Qu.: 95.0   1st Qu.: 7.400
 Median : 21.00   Median :194.0   Median : 9.700
 Mean   : 31.94   Mean   :177.4   Mean   : 9.958
 3rd Qu.: 46.00   3rd Qu.:256.0   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
      Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0
```

Replace the NA values in the attribute Ozone in aq by mean of the remaining values. Print the summary of the dataset aq

```
> summary(airquality$Ozone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   1.00   18.00   31.50   42.13   63.25  168.00      37

> aq$Ozone[is.na(aq$Ozone)] <- mean(airquality$Ozone, na.rm = TRUE)
>

> summary(aq$Ozone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00    4.00   21.00   31.94   46.00  168.00

> summary(aq)
     Ozone           Solar.R          Wind
 Min.   :  0.00   Min.   :  0.0   Min.   : 1.700
 1st Qu.:  4.00   1st Qu.: 95.0   1st Qu.: 7.400
 Median : 21.00   Median :194.0   Median : 9.700
 Mean   : 31.94   Mean   :177.4   Mean   : 9.958
 3rd Qu.: 46.00   3rd Qu.:256.0   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
      Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0
```

Copy the dataset airquality to aq1. Replace the NA values in the attribute Ozone in aq1 by median of the remaining values. Print the summary of the dataset aq1

```
> aq1 <- airquality
>
```

```
> aq1$Ozone[is.na(aq$Ozone)] <- median(aq1$Ozone, na.rm = TRUE)
>
```

```
> summary(aq1$Ozone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   1.00   18.00   31.50   42.13   63.25  168.00      37
```

```
> aq1$Ozone[is.na(aq$Ozone)] <- median(aq1$Ozone, na.rm = TRUE)
>
> summary(aq1$Ozone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   1.00   18.00   31.50   42.13   63.25  168.00      37
```

```
> aq1$Ozone[is.na(aq1$Ozone)] <- median(aq1$Ozone, na.rm = TRUE)
> summary(aq1$Ozone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1.00   21.00   31.50   39.56   46.00  168.00
```

Copy the dataset airquality to aq2. Replace the NA values in the attribute Ozone in aq2 by mode of the remaining values. Print the summary of the dataset aq2

```
> aq2 <- airquality
>
```

Repeat the above five operations for the attribute Solar.R

Replace all the values of Temp with global constant 50 in aq1

Replace all the values below 60 of Temp with global constant 60 in aq2

Replace the month numbers in the column Month in aq by name of the month. (Ex: Replace 5

with May). (Hint: use gsub() function. aq$Month <- gsub(5,"May",aq$Month))

Create a new logical attribute Solar.Danger in aq by filling it's value with TRUE if the value in the attribute Solar.R is greater than 100, other with FALSE

Discretize the values in Temp of aq to "Low", "Medium" and "High"

What does cut() function do?

Create a numeric vector brks containing values 0, 50, 100, 200, 250, 300 and 350. Divide the range of Solar.R into intervals and recode the values in Solar.R according to which interval they fall using the vector brks.

aq$Solar.R=cut(aq$Solar.R,breaks=brks,include.lowest=TRUE)

Practice the examples given in lecture slide 12 on dataset airquality

52	NA	150	6.3	77	6	21
53	NA	59	1.7	76	6	22
54	NA	91	4.6	76	6	23
55	NA	250	6.3	76	6	24
56	NA	135	8.0	75	6	25
57	NA	127	8.0	78	6	26
58	NA	47	10.3	73	6	27
59	NA	98	11.5	80	6	28
60	NA	31	14.9	77	6	29
61	NA	138	8.0	83	6	30
62	135	269	4.1	84	7	1
63	49	248	9.2	85	7	2

64	32	236	9.2	81	7	3
65	NA	101	10.9	84	7	4
66	64	175	4.6	83	7	5
67	40	314	10.9	83	7	6
68	77	276	5.1	88	7	7
69	97	267	6.3	92	7	8
70	97	272	5.7	92	7	9
71	85	175	7.4	89	7	10
72	NA	139	8.6	82	7	11
73	10	264	14.3	73	7	12
74	27	175	14.9	81	7	13
75	NA	291	14.9	91	7	14
76	7	48	14.3	80	7	15
77	48	260	6.9	81	7	16
78	35	274	10.3	82	7	17
79	61	285	6.3	84	7	18
80	79	187	5.1	87	7	19
81	63	220	11.5	85	7	20
82	16	7	6.9	74	7	21
83	NA	258	9.7	81	7	22
84	NA	295	11.5	82	7	23
85	80	294	8.6	86	7	24
86	108	223	8.0	85	7	25
87	20	81	8.6	82	7	26
88	52	82	12.0	86	7	27
89	82	213	7.4	88	7	28
90	50	275	7.4	86	7	29
91	64	253	7.4	83	7	30
92	59	254	9.2	81	7	31
93	39	83	6.9	81	8	1
94	9	24	13.8	81	8	2
95	16	77	7.4	82	8	3
96	78	NA	6.9	86	8	4
97	35	NA	7.4	85	8	5
98	66	NA	4.6	87	8	6
99	122	255	4.0	89	8	7
100	89	229	10.3	90	8	8
101	110	207	8.0	90	8	9
102	NA	222	8.6	92	8	10
103	NA	137	11.5	86	8	11
104	44	192	11.5	86	8	12
105	28	273	11.5	82	8	13
106	65	157	9.7	80	8	14
107	NA	64	11.5	79	8	15
108	22	71	10.3	77	8	16
109	59	51	6.3	79	8	17
110	23	115	7.4	76	8	18
111	31	244	10.9	78	8	19

```
112    44     190 10.3   78     8  20
113    21     259 15.5   77     8  21
114     9      36 14.3   72     8  22
115    NA     255 12.6   75     8  23
116    45     212  9.7   79     8  24
117   168     238  3.4   81     8  25
118    73     215  8.0   86     8  26
119    NA     153  5.7   88     8  27
120    76     203  9.7   97     8  28
121   118     225  2.3   94     8  29
122    84     237  6.3   96     8  30
123    85     188  6.3   94     8  31
124    96     167  6.9   91     9   1
125    78     197  5.1   92     9   2
126    73     183  2.8   93     9   3
127    91     189  4.6   93     9   4
128    47      95  7.4   87     9   5
129    32      92 15.5   84     9   6
130    20     252 10.9   80     9   7
131    23     220 10.3   78     9   8
132    21     230 10.9   75     9   9
133    24     259  9.7   73     9  10
134    44     236 14.9   81     9  11
135    21     259 15.5   76     9  12
136    28     238  6.3   77     9  13
137     9      24 10.9   71     9  14
138    13     112 11.5   71     9  15
139    46     237  6.9   78     9  16
140    18     224 13.8   67     9  17
141    13      27 10.3   76     9  18
142    24     238 10.3   68     9  19
143    16     201  8.0   82     9  20
144    13     238 12.6   64     9  21
145    23      14  9.2   71     9  22
146    36     139 10.3   81     9  23
147     7      49 10.3   69     9  24
148    14      20 16.6   63     9  25
149    30     193  6.9   70     9  26
150    NA     145 13.2   77     9  27
151    14     191 14.3   75     9  28
152    18     131  8.0   76     9  29
153    20     223 11.5   68     9  30
> str(airquality)
'data.frame':    153 obs. of  6 variables:
 $ Ozone  : int  41 36 12 18 NA 28 23 19 8 NA ...
 $ Solar.R: int  190 118 149 313 NA NA 299 99 19 194 ...
 $ Wind   : num  7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6 ...
 $ Temp   : int  67 72 74 62 56 66 65 59 61 69 ...
```

```
 $ Month  : int  5 5 5 5 5 5 5 5 5 5 ...
 $ Day    : int  1 2 3 4 5 6 7 8 9 10 ...
>
>
> summary()
Error in is.factor(object) :
  argument "object" is missing, with no default
> summary(airquality)
     Ozone           Solar.R          Wind
 Min.   :  1.00   Min.   :  7.0   Min.   : 1.700
 1st Qu.: 18.00   1st Qu.:115.8   1st Qu.: 7.400
 Median : 31.50   Median :205.0   Median : 9.700
 Mean   : 42.13   Mean   :185.9   Mean   : 9.958
 3rd Qu.: 63.25   3rd Qu.:258.8   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
 NA's   :37       NA's   :7
     Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0

>
>
>
>
>
> summary(airquality)
     Ozone           Solar.R          Wind
 Min.   :  1.00   Min.   :  7.0   Min.   : 1.700
 1st Qu.: 18.00   1st Qu.:115.8   1st Qu.: 7.400
 Median : 31.50   Median :205.0   Median : 9.700
 Mean   : 42.13   Mean   :185.9   Mean   : 9.958
 3rd Qu.: 63.25   3rd Qu.:258.8   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
 NA's   :37       NA's   :7
     Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0

>
>
```

```
>
> ncol(airquality)
[1] 6
>
> nrow(airquality)
[1] 153
>
>
>
>
> head(airqua;ity)
Error: unexpected ';' in "head(airqua;"
> head(airquaity)
Error in head(airquaity) : object 'airquaity' not found
> head(airquality)
  Ozone Solar.R Wind Temp Month Day
1    41     190  7.4   67     5   1
2    36     118  8.0   72     5   2
3    12     149 12.6   74     5   3
4    18     313 11.5   62     5   4
5    NA      NA 14.3   56     5   5
6    28      NA 14.9   66     5   6
>
>
>
>
> head(airquality)
  Ozone Solar.R Wind Temp Month Day
1    41     190  7.4   67     5   1
2    36     118  8.0   72     5   2
3    12     149 12.6   74     5   3
4    18     313 11.5   62     5   4
5    NA      NA 14.3   56     5   5
6    28      NA 14.9   66     5   6
>
> x[x<6]=0
>
>
>
>
>
>
>
> x
[1] 0 0 0 6 7 8
>
>
>
```

```
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
> x[x<6]=0
> x
[1] 0 0 0 6 7 8
```

```
>
> x[x>6 & x<8]=100
>
> x
[1]   0   0   0   6 100   8
>
> x[x<5 | x>6]<-50
> x
[1] 50 50 50  6 50 50
>
> x[x<6]<-NA
> x
[1] 50 50 50  6 50 50
>
> x[x>4 & x<8] <-NA
> x
[1] 50 50 50 NA 50 50
>
> x[x<5 | x>6]<-NA
> x
[1] NA NA NA NA NA NA
>
> sum(is.na(x))
[1] 6
>
> mean(x, na.rm=TRUE)
[1] NaN
>
>
> median(x, na.rm=TRUE)
[1] NA
>
> x[x<6]<-"NA"
> x
[1] NA NA NA NA NA NA
> x=as.integer(x)
>
> x
[1] NA NA NA NA NA NA
> x=as.integer(x)
> x
[1] NA NA NA NA NA NA
>
>
> x[x<6]<-"NA"
> x
[1] NA NA NA NA NA NA
> x=as.integer(x)
```

```
> x
[1] NA NA NA NA NA NA
> A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
>
>
>
>
>
>
>
> A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
> B <- c(3, 2, NA, 5, 3, 7, NA, “NA”, 5, 2, 6)
Error: unexpected input in "B <- c(3, 2, NA, 5, 3, 7, NA, “"
>
> length(A)
[1] 11
> length(B)
Error: object 'B' not found
> B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)
>
>
>
>
>
>
>
>
> A <- c(3, 2, NA, 5, 3, 7, NA, NA, 5, 2, 6)
> B <- c(3, 2, NA, 5, 3, 7, NA, "NA", 5, 2, 6)
> length(A)
[1] 11
> length(B)
[1] 11
>
>
> p <- sort(A)
>
> p
[1] 2 2 3 3 5 5 6 7
>
>
> p <- sort(A)
> p
[1] 2 2 3 3 5 5 6 7
>
> length(p)
[1] 8
>
```

```
> q <- sort(B)
> q
[1] "2"  "2"  "3"  "3"  "5"  "5"  "6"  "7"  "NA"
>
> length(q)
[1] 9
>
> x <- c(3,4,5,6,7,8)
>
> x=as.integer(x)
> x
[1] 3 4 5 6 7 8
>
> x[x<6] <- "NA"
> x
[1] "NA" "NA" "NA" "6"  "7"  "8"
> x=as.integer(x)
Warning message:
NAs introduced by coercion
>
>
> airquality {datasets}
Error: unexpected '{' in "airquality {"
>
>
> aq <- airquality
>
> aq
    Ozone Solar.R Wind Temp Month Day
1      41     190  7.4   67     5   1
2      36     118  8.0   72     5   2
3      12     149 12.6   74     5   3
4      18     313 11.5   62     5   4
5      NA      NA 14.3   56     5   5
6      28      NA 14.9   66     5   6
7      23     299  8.6   65     5   7
8      19      99 13.8   59     5   8
9       8      19 20.1   61     5   9
10     NA     194  8.6   69     5  10
11      7      NA  6.9   74     5  11
12     16     256  9.7   69     5  12
13     11     290  9.2   66     5  13
14     14     274 10.9   68     5  14
15     18      65 13.2   58     5  15
16     14     334 11.5   64     5  16
17     34     307 12.0   66     5  17
18      6      78 18.4   57     5  18
19     30     322 11.5   68     5  19
```

20	11	44	9.7	62	5	20
21	1	8	9.7	59	5	21
22	11	320	16.6	73	5	22
23	4	25	9.7	61	5	23
24	32	92	12.0	61	5	24
25	NA	66	16.6	57	5	25
26	NA	266	14.9	58	5	26
27	NA	NA	8.0	57	5	27
28	23	13	12.0	67	5	28
29	45	252	14.9	81	5	29
30	115	223	5.7	79	5	30
31	37	279	7.4	76	5	31
32	NA	286	8.6	78	6	1
33	NA	287	9.7	74	6	2
34	NA	242	16.1	67	6	3
35	NA	186	9.2	84	6	4
36	NA	220	8.6	85	6	5
37	NA	264	14.3	79	6	6
38	29	127	9.7	82	6	7
39	NA	273	6.9	87	6	8
40	71	291	13.8	90	6	9
41	39	323	11.5	87	6	10
42	NA	259	10.9	93	6	11
43	NA	250	9.2	92	6	12
44	23	148	8.0	82	6	13
45	NA	332	13.8	80	6	14
46	NA	322	11.5	79	6	15
47	21	191	14.9	77	6	16
48	37	284	20.7	72	6	17
49	20	37	9.2	65	6	18
50	12	120	11.5	73	6	19
51	13	137	10.3	76	6	20
52	NA	150	6.3	77	6	21
53	NA	59	1.7	76	6	22
54	NA	91	4.6	76	6	23
55	NA	250	6.3	76	6	24
56	NA	135	8.0	75	6	25
57	NA	127	8.0	78	6	26
58	NA	47	10.3	73	6	27
59	NA	98	11.5	80	6	28
60	NA	31	14.9	77	6	29
61	NA	138	8.0	83	6	30
62	135	269	4.1	84	7	1
63	49	248	9.2	85	7	2
64	32	236	9.2	81	7	3
65	NA	101	10.9	84	7	4
66	64	175	4.6	83	7	5
67	40	314	10.9	83	7	6

68	77	276	5.1	88	7	7
69	97	267	6.3	92	7	8
70	97	272	5.7	92	7	9
71	85	175	7.4	89	7	10
72	NA	139	8.6	82	7	11
73	10	264	14.3	73	7	12
74	27	175	14.9	81	7	13
75	NA	291	14.9	91	7	14
76	7	48	14.3	80	7	15
77	48	260	6.9	81	7	16
78	35	274	10.3	82	7	17
79	61	285	6.3	84	7	18
80	79	187	5.1	87	7	19
81	63	220	11.5	85	7	20
82	16	7	6.9	74	7	21
83	NA	258	9.7	81	7	22
84	NA	295	11.5	82	7	23
85	80	294	8.6	86	7	24
86	108	223	8.0	85	7	25
87	20	81	8.6	82	7	26
88	52	82	12.0	86	7	27
89	82	213	7.4	88	7	28
90	50	275	7.4	86	7	29
91	64	253	7.4	83	7	30
92	59	254	9.2	81	7	31
93	39	83	6.9	81	8	1
94	9	24	13.8	81	8	2
95	16	77	7.4	82	8	3
96	78	NA	6.9	86	8	4
97	35	NA	7.4	85	8	5
98	66	NA	4.6	87	8	6
99	122	255	4.0	89	8	7
100	89	229	10.3	90	8	8
101	110	207	8.0	90	8	9
102	NA	222	8.6	92	8	10
103	NA	137	11.5	86	8	11
104	44	192	11.5	86	8	12
105	28	273	11.5	82	8	13
106	65	157	9.7	80	8	14
107	NA	64	11.5	79	8	15
108	22	71	10.3	77	8	16
109	59	51	6.3	79	8	17
110	23	115	7.4	76	8	18
111	31	244	10.9	78	8	19
112	44	190	10.3	78	8	20
113	21	259	15.5	77	8	21
114	9	36	14.3	72	8	22
115	NA	255	12.6	75	8	23

```
116    45    212  9.7   79     8  24
117   168    238  3.4   81     8  25
118    73    215  8.0   86     8  26
119    NA    153  5.7   88     8  27
120    76    203  9.7   97     8  28
121   118    225  2.3   94     8  29
122    84    237  6.3   96     8  30
123    85    188  6.3   94     8  31
124    96    167  6.9   91     9   1
125    78    197  5.1   92     9   2
126    73    183  2.8   93     9   3
127    91    189  4.6   93     9   4
128    47     95  7.4   87     9   5
129    32     92 15.5   84     9   6
130    20    252 10.9   80     9   7
131    23    220 10.3   78     9   8
132    21    230 10.9   75     9   9
133    24    259  9.7   73     9  10
134    44    236 14.9   81     9  11
135    21    259 15.5   76     9  12
136    28    238  6.3   77     9  13
137     9     24 10.9   71     9  14
138    13    112 11.5   71     9  15
139    46    237  6.9   78     9  16
140    18    224 13.8   67     9  17
141    13     27 10.3   76     9  18
142    24    238 10.3   68     9  19
143    16    201  8.0   82     9  20
144    13    238 12.6   64     9  21
145    23     14  9.2   71     9  22
146    36    139 10.3   81     9  23
147     7     49 10.3   69     9  24
148    14     20 16.6   63     9  25
149    30    193  6.9   70     9  26
150    NA    145 13.2   77     9  27
151    14    191 14.3   75     9  28
152    18    131  8.0   76     9  29
153    20    223 11.5   68     9  30
> print(aq)
    Ozone Solar.R Wind Temp Month Day
1      41     190  7.4   67     5   1
2      36     118  8.0   72     5   2
3      12     149 12.6   74     5   3
4      18     313 11.5   62     5   4
5      NA      NA 14.3   56     5   5
6      28      NA 14.9   66     5   6
7      23     299  8.6   65     5   7
8      19      99 13.8   59     5   8
```

9	8	19	20.1	61	5	9
10	NA	194	8.6	69	5	10
11	7	NA	6.9	74	5	11
12	16	256	9.7	69	5	12
13	11	290	9.2	66	5	13
14	14	274	10.9	68	5	14
15	18	65	13.2	58	5	15
16	14	334	11.5	64	5	16
17	34	307	12.0	66	5	17
18	6	78	18.4	57	5	18
19	30	322	11.5	68	5	19
20	11	44	9.7	62	5	20
21	1	8	9.7	59	5	21
22	11	320	16.6	73	5	22
23	4	25	9.7	61	5	23
24	32	92	12.0	61	5	24
25	NA	66	16.6	57	5	25
26	NA	266	14.9	58	5	26
27	NA	NA	8.0	57	5	27
28	23	13	12.0	67	5	28
29	45	252	14.9	81	5	29
30	115	223	5.7	79	5	30
31	37	279	7.4	76	5	31
32	NA	286	8.6	78	6	1
33	NA	287	9.7	74	6	2
34	NA	242	16.1	67	6	3
35	NA	186	9.2	84	6	4
36	NA	220	8.6	85	6	5
37	NA	264	14.3	79	6	6
38	29	127	9.7	82	6	7
39	NA	273	6.9	87	6	8
40	71	291	13.8	90	6	9
41	39	323	11.5	87	6	10
42	NA	259	10.9	93	6	11
43	NA	250	9.2	92	6	12
44	23	148	8.0	82	6	13
45	NA	332	13.8	80	6	14
46	NA	322	11.5	79	6	15
47	21	191	14.9	77	6	16
48	37	284	20.7	72	6	17
49	20	37	9.2	65	6	18
50	12	120	11.5	73	6	19
51	13	137	10.3	76	6	20
52	NA	150	6.3	77	6	21
53	NA	59	1.7	76	6	22
54	NA	91	4.6	76	6	23
55	NA	250	6.3	76	6	24
56	NA	135	8.0	75	6	25

57	NA	127	8.0	78	6	26
58	NA	47	10.3	73	6	27
59	NA	98	11.5	80	6	28
60	NA	31	14.9	77	6	29
61	NA	138	8.0	83	6	30
62	135	269	4.1	84	7	1
63	49	248	9.2	85	7	2
64	32	236	9.2	81	7	3
65	NA	101	10.9	84	7	4
66	64	175	4.6	83	7	5
67	40	314	10.9	83	7	6
68	77	276	5.1	88	7	7
69	97	267	6.3	92	7	8
70	97	272	5.7	92	7	9
71	85	175	7.4	89	7	10
72	NA	139	8.6	82	7	11
73	10	264	14.3	73	7	12
74	27	175	14.9	81	7	13
75	NA	291	14.9	91	7	14
76	7	48	14.3	80	7	15
77	48	260	6.9	81	7	16
78	35	274	10.3	82	7	17
79	61	285	6.3	84	7	18
80	79	187	5.1	87	7	19
81	63	220	11.5	85	7	20
82	16	7	6.9	74	7	21
83	NA	258	9.7	81	7	22
84	NA	295	11.5	82	7	23
85	80	294	8.6	86	7	24
86	108	223	8.0	85	7	25
87	20	81	8.6	82	7	26
88	52	82	12.0	86	7	27
89	82	213	7.4	88	7	28
90	50	275	7.4	86	7	29
91	64	253	7.4	83	7	30
92	59	254	9.2	81	7	31
93	39	83	6.9	81	8	1
94	9	24	13.8	81	8	2
95	16	77	7.4	82	8	3
96	78	NA	6.9	86	8	4
97	35	NA	7.4	85	8	5
98	66	NA	4.6	87	8	6
99	122	255	4.0	89	8	7
100	89	229	10.3	90	8	8
101	110	207	8.0	90	8	9
102	NA	222	8.6	92	8	10
103	NA	137	11.5	86	8	11
104	44	192	11.5	86	8	12

105	28	273	11.5	82	8	13
106	65	157	9.7	80	8	14
107	NA	64	11.5	79	8	15
108	22	71	10.3	77	8	16
109	59	51	6.3	79	8	17
110	23	115	7.4	76	8	18
111	31	244	10.9	78	8	19
112	44	190	10.3	78	8	20
113	21	259	15.5	77	8	21
114	9	36	14.3	72	8	22
115	NA	255	12.6	75	8	23
116	45	212	9.7	79	8	24
117	168	238	3.4	81	8	25
118	73	215	8.0	86	8	26
119	NA	153	5.7	88	8	27
120	76	203	9.7	97	8	28
121	118	225	2.3	94	8	29
122	84	237	6.3	96	8	30
123	85	188	6.3	94	8	31
124	96	167	6.9	91	9	1
125	78	197	5.1	92	9	2
126	73	183	2.8	93	9	3
127	91	189	4.6	93	9	4
128	47	95	7.4	87	9	5
129	32	92	15.5	84	9	6
130	20	252	10.9	80	9	7
131	23	220	10.3	78	9	8
132	21	230	10.9	75	9	9
133	24	259	9.7	73	9	10
134	44	236	14.9	81	9	11
135	21	259	15.5	76	9	12
136	28	238	6.3	77	9	13
137	9	24	10.9	71	9	14
138	13	112	11.5	71	9	15
139	46	237	6.9	78	9	16
140	18	224	13.8	67	9	17
141	13	27	10.3	76	9	18
142	24	238	10.3	68	9	19
143	16	201	8.0	82	9	20
144	13	238	12.6	64	9	21
145	23	14	9.2	71	9	22
146	36	139	10.3	81	9	23
147	7	49	10.3	69	9	24
148	14	20	16.6	63	9	25
149	30	193	6.9	70	9	26
150	NA	145	13.2	77	9	27
151	14	191	14.3	75	9	28
152	18	131	8.0	76	9	29

```
153    20     223 11.5   68     9  30
>
>
> str(aq)
'data.frame':    153 obs. of  6 variables:
 $ Ozone  : int  41 36 12 18 NA 28 23 19 8 NA ...
 $ Solar.R: int  190 118 149 313 NA NA 299 99 19 194 ...
 $ Wind   : num  7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6 ...
 $ Temp   : int  67 72 74 62 56 66 65 59 61 69 ...
 $ Month  : int  5 5 5 5 5 5 5 5 5 5 ...
 $ Day    : int  1 2 3 4 5 6 7 8 9 10 ...
>
>
> summary(aq)
     Ozone           Solar.R          Wind
 Min.   :  1.00   Min.   :  7.0   Min.   : 1.700
 1st Qu.: 18.00   1st Qu.:115.8   1st Qu.: 7.400
 Median : 31.50   Median :205.0   Median : 9.700
 Mean   : 42.13   Mean   :185.9   Mean   : 9.958
 3rd Qu.: 63.25   3rd Qu.:258.8   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
 NA's   :37       NA's   :7
      Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0

>
> tail(aq)
    Ozone Solar.R Wind Temp Month Day
148    14      20 16.6   63     9  25
149    30     193  6.9   70     9  26
150    NA     145 13.2   77     9  27
151    14     191 14.3   75     9  28
152    18     131  8.0   76     9  29
153    20     223 11.5   68     9  30
>
>
> head(aq)
  Ozone Solar.R Wind Temp Month Day
1    41     190  7.4   67     5   1
2    36     118  8.0   72     5   2
3    12     149 12.6   74     5   3
4    18     313 11.5   62     5   4
5    NA      NA 14.3   56     5   5
```

```
6   28      NA 14.9   66     5   6
>
>
> aq[is.na(aq)] <- 0
>
> summary(aq)
     Ozone           Solar.R          Wind
 Min.   :  0.00   Min.   :  0.0   Min.   : 1.700
 1st Qu.:  4.00   1st Qu.: 95.0   1st Qu.: 7.400
 Median : 21.00   Median :194.0   Median : 9.700
 Mean   : 31.94   Mean   :177.4   Mean   : 9.958
 3rd Qu.: 46.00   3rd Qu.:256.0   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
      Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0
>
> x <- c(3,4,5,6,7,8)
> x[x>4 & x<8] <-NA
> x <- c(3,4,5,6,7,8)
> x[x>4 & x<8] <- "NA"
> x
[1] "3"  "4"  "NA" "NA" "NA" "8"
>
> x[x<5 | x>6]<- "NA"
> X
Error: object 'X' not found
>
>
> x[x<5 | x>6]<- "NA"
> x
[1] "NA" "NA" "NA" "NA" "NA" "NA"
>
> x <- c(3,4,5,6,7,8)
> x[x<5 | x>6]<- "NA"
> x
[1] "NA" "NA" "5"  "6"  "NA" "NA"
>
> mean(x, na.rm=TRUE)
[1] NA
Warning message:
In mean.default(x, na.rm = TRUE) :
  argument is not numeric or logical: returning NA
>
```

```
> x=as.integer(x)
Warning message:
NAs introduced by coercion
>
>
> colSums(is.na(aq))
  Ozone Solar.R    Wind    Temp   Month     Day
      0       0       0       0       0       0
>
> colSums(is.na(airquality))
  Ozone Solar.R    Wind    Temp   Month     Day
     37       7       0       0       0       0
>
> datasets::airquality
    Ozone Solar.R Wind Temp Month Day
1      41     190  7.4   67     5   1
2      36     118  8.0   72     5   2
3      12     149 12.6   74     5   3
4      18     313 11.5   62     5   4
5      NA      NA 14.3   56     5   5
6      28      NA 14.9   66     5   6
7      23     299  8.6   65     5   7
8      19      99 13.8   59     5   8
9       8      19 20.1   61     5   9
10     NA     194  8.6   69     5  10
11      7      NA  6.9   74     5  11
12     16     256  9.7   69     5  12
13     11     290  9.2   66     5  13
14     14     274 10.9   68     5  14
15     18      65 13.2   58     5  15
16     14     334 11.5   64     5  16
17     34     307 12.0   66     5  17
18      6      78 18.4   57     5  18
19     30     322 11.5   68     5  19
20     11      44  9.7   62     5  20
21      1       8  9.7   59     5  21
22     11     320 16.6   73     5  22
23      4      25  9.7   61     5  23
24     32      92 12.0   61     5  24
25     NA      66 16.6   57     5  25
26     NA     266 14.9   58     5  26
27     NA      NA  8.0   57     5  27
28     23      13 12.0   67     5  28
29     45     252 14.9   81     5  29
30    115     223  5.7   79     5  30
31     37     279  7.4   76     5  31
32     NA     286  8.6   78     6   1
33     NA     287  9.7   74     6   2
```

34	NA	242	16.1	67	6	3
35	NA	186	9.2	84	6	4
36	NA	220	8.6	85	6	5
37	NA	264	14.3	79	6	6
38	29	127	9.7	82	6	7
39	NA	273	6.9	87	6	8
40	71	291	13.8	90	6	9
41	39	323	11.5	87	6	10
42	NA	259	10.9	93	6	11
43	NA	250	9.2	92	6	12
44	23	148	8.0	82	6	13
45	NA	332	13.8	80	6	14
46	NA	322	11.5	79	6	15
47	21	191	14.9	77	6	16
48	37	284	20.7	72	6	17
49	20	37	9.2	65	6	18
50	12	120	11.5	73	6	19
51	13	137	10.3	76	6	20
52	NA	150	6.3	77	6	21
53	NA	59	1.7	76	6	22
54	NA	91	4.6	76	6	23
55	NA	250	6.3	76	6	24
56	NA	135	8.0	75	6	25
57	NA	127	8.0	78	6	26
58	NA	47	10.3	73	6	27
59	NA	98	11.5	80	6	28
60	NA	31	14.9	77	6	29
61	NA	138	8.0	83	6	30
62	135	269	4.1	84	7	1
63	49	248	9.2	85	7	2
64	32	236	9.2	81	7	3
65	NA	101	10.9	84	7	4
66	64	175	4.6	83	7	5
67	40	314	10.9	83	7	6
68	77	276	5.1	88	7	7
69	97	267	6.3	92	7	8
70	97	272	5.7	92	7	9
71	85	175	7.4	89	7	10
72	NA	139	8.6	82	7	11
73	10	264	14.3	73	7	12
74	27	175	14.9	81	7	13
75	NA	291	14.9	91	7	14
76	7	48	14.3	80	7	15
77	48	260	6.9	81	7	16
78	35	274	10.3	82	7	17
79	61	285	6.3	84	7	18
80	79	187	5.1	87	7	19
81	63	220	11.5	85	7	20

82	16	7	6.9	74	7	21
83	NA	258	9.7	81	7	22
84	NA	295	11.5	82	7	23
85	80	294	8.6	86	7	24
86	108	223	8.0	85	7	25
87	20	81	8.6	82	7	26
88	52	82	12.0	86	7	27
89	82	213	7.4	88	7	28
90	50	275	7.4	86	7	29
91	64	253	7.4	83	7	30
92	59	254	9.2	81	7	31
93	39	83	6.9	81	8	1
94	9	24	13.8	81	8	2
95	16	77	7.4	82	8	3
96	78	NA	6.9	86	8	4
97	35	NA	7.4	85	8	5
98	66	NA	4.6	87	8	6
99	122	255	4.0	89	8	7
100	89	229	10.3	90	8	8
101	110	207	8.0	90	8	9
102	NA	222	8.6	92	8	10
103	NA	137	11.5	86	8	11
104	44	192	11.5	86	8	12
105	28	273	11.5	82	8	13
106	65	157	9.7	80	8	14
107	NA	64	11.5	79	8	15
108	22	71	10.3	77	8	16
109	59	51	6.3	79	8	17
110	23	115	7.4	76	8	18
111	31	244	10.9	78	8	19
112	44	190	10.3	78	8	20
113	21	259	15.5	77	8	21
114	9	36	14.3	72	8	22
115	NA	255	12.6	75	8	23
116	45	212	9.7	79	8	24
117	168	238	3.4	81	8	25
118	73	215	8.0	86	8	26
119	NA	153	5.7	88	8	27
120	76	203	9.7	97	8	28
121	118	225	2.3	94	8	29
122	84	237	6.3	96	8	30
123	85	188	6.3	94	8	31
124	96	167	6.9	91	9	1
125	78	197	5.1	92	9	2
126	73	183	2.8	93	9	3
127	91	189	4.6	93	9	4
128	47	95	7.4	87	9	5
129	32	92	15.5	84	9	6

```
130    20     252 10.9   80     9   7
131    23     220 10.3   78     9   8
132    21     230 10.9   75     9   9
133    24     259  9.7   73     9  10
134    44     236 14.9   81     9  11
135    21     259 15.5   76     9  12
136    28     238  6.3   77     9  13
137     9      24 10.9   71     9  14
138    13     112 11.5   71     9  15
139    46     237  6.9   78     9  16
140    18     224 13.8   67     9  17
141    13      27 10.3   76     9  18
142    24     238 10.3   68     9  19
143    16     201  8.0   82     9  20
144    13     238 12.6   64     9  21
145    23      14  9.2   71     9  22
146    36     139 10.3   81     9  23
147     7      49 10.3   69     9  24
148    14      20 16.6   63     9  25
149    30     193  6.9   70     9  26
150    NA     145 13.2   77     9  27
151    14     191 14.3   75     9  28
152    18     131  8.0   76     9  29
153    20     223 11.5   68     9  30
>
> data(airquality)
>
> ?airquality
>
>
> aq$Solar.R=cut(aq$Solar.R,breaks=brks,include.lowest=TRUE)
Error in cut.default(aq$Solar.R, breaks = brks, include.lowest = TRUE) :
  object 'brks' not found
>
>
>
>
> mean(airquality$Ozone, na.rm = TRUE)
[1] 42.12931
>
> airquality[is.na(airquality$Ozone)]
Error in `[.data.frame`(airquality, is.na(airquality$Ozone)) :
  undefined columns selected
>
> airquality$Ozone[is.na(airquality$Ozone)]
 [1] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
[18] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
[35] NA NA NA
```

```
>
> aq$Ozone[is.na(aq$Ozone)] <- mean(airquality$Ozone, na.rm = TRUE)
>
> summary(aq)
     Ozone           Solar.R          Wind
 Min.   :  0.00   Min.   :  0.0   Min.   : 1.700
 1st Qu.:  4.00   1st Qu.: 95.0   1st Qu.: 7.400
 Median : 21.00   Median :194.0   Median : 9.700
 Mean   : 31.94   Mean   :177.4   Mean   : 9.958
 3rd Qu.: 46.00   3rd Qu.:256.0   3rd Qu.:11.500
 Max.   :168.00   Max.   :334.0   Max.   :20.700
      Temp           Month            Day
 Min.   :56.00   Min.   :5.000   Min.   : 1.0
 1st Qu.:72.00   1st Qu.:6.000   1st Qu.: 8.0
 Median :79.00   Median :7.000   Median :16.0
 Mean   :77.88   Mean   :6.993   Mean   :15.8
 3rd Qu.:85.00   3rd Qu.:8.000   3rd Qu.:23.0
 Max.   :97.00   Max.   :9.000   Max.   :31.0
>
> summary(aq$Ozone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00    4.00   21.00   31.94   46.00  168.00
>
> summary(airquality$Ozone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   1.00   18.00   31.50   42.13   63.25  168.00      37
```